Supporting Information

A New Mass Spectrometry-compatible Degradable Surfactant for Tissue Proteomics

Ying-Hua Chang,[1] Zachery R. Gregorich,[1,2] Albert J. Chen,[1] Leekyoung Hwang,[3] Huseyin Guner,[4] Deyang Yu,[1,5] Jianyi Zhang,[6] Ying Ge[1-5*]

[1]Department of Cell and Regenerative Biology, [2]Molecular and Cellular Pharmacology Program, [3]Department of Chemistry, [4]Human Proteomics Program, [5]Molecular and Environmental Toxicology Program, University of Wisconsin-Madison, Madison, WI, USA

[6]Division of Cardiology, Department of Medicine, Department of Biomedical Engineering, Stem Cell Institute, University of Minnesota, Minneapolis, MN, 55455, USA

*Correspondence:

Prof. Ying Ge

Department of Cell and Regenerative Biology

Department of Chemistry

University of Wisconsin-Madison

1300 University Ave., SMI 130, Madison, WI 53706

Tel: 608-263-9212, Fax: 608-265-5512, E-mail: ge2@wisc.edu.

Supplemental Method

MS analysis of BSA (LTQ)

MaSDeS and SDS were spiked into samples containing bovine serum albumin (BSA) so that the final concentration of surfactant was 0.1%, 0.2%, or 0.5%. In-solution digestion was then performed as described above. A reverse phase column attached to an Eksigent NanoLC-2D (Eksigent, CA, USA) was used to separate the digested peptide mixtures. When performing the nanoLC-ESI-MS/MS experiments, a peptide nanotrap column (300 µm x 5 mm, C18, 3 µm, 100 Å) (Dionex Corporation, CA, USA) along with a home-made analytical column (75 µm x 100 mm, C18) were set in place prior to the electrospray source. The peptide sample was injected and cleaned on the trap column with 5% solvent B (0.1% formic acid in 100% acetonitrile) at 10 µL/min for 10 min. Thereafter, the peptides were eluted using a 35 min gradient from 5% to 35% B at a flow rate of 500 nL/min (where solvent A was formed by 0.1% formic acid in HPLC grade water). An LTQ mass spectrometer (Thermo Scientific, Bremen, Germany) was used to acquire spectra, and was operated in data-dependent mode with dynamic exclusion enabled. The MS/MS spectra were obtained for the five most abundant peptide ions in full MS scan and were then searched against the bovine protein database from SwissProt using the SEQUEST algorithm found in the Bioworks software. The precursor mass and fragment ion tolerances were set as 1 Da and 1 Da, respectively. Deamidated asparagine and glutamine with oxidized methionine were used as variable modifications. Carbamidomethyl of cysteine served as a fixed modification during the database search. Trypsin was used as a specific protease with two missed cleavages allowed. The filters set for peptides identification were delta CN 0.1, Rep 5, Xcorr vs charge states of 1.9, 2.2, and 3.75 for charges 1, 2, and 3, respectively, and a probability of 0.001. For protein identification, the filters were assigned a protein probability of 0.0001 and required 2

unique peptides. The peptide peak areas were quantified as they passed the minimal threshold of 50,000 and a smoothing point set at 5.

Supplemental Results

Improvement of tissue protein identification using sequential tissue extraction

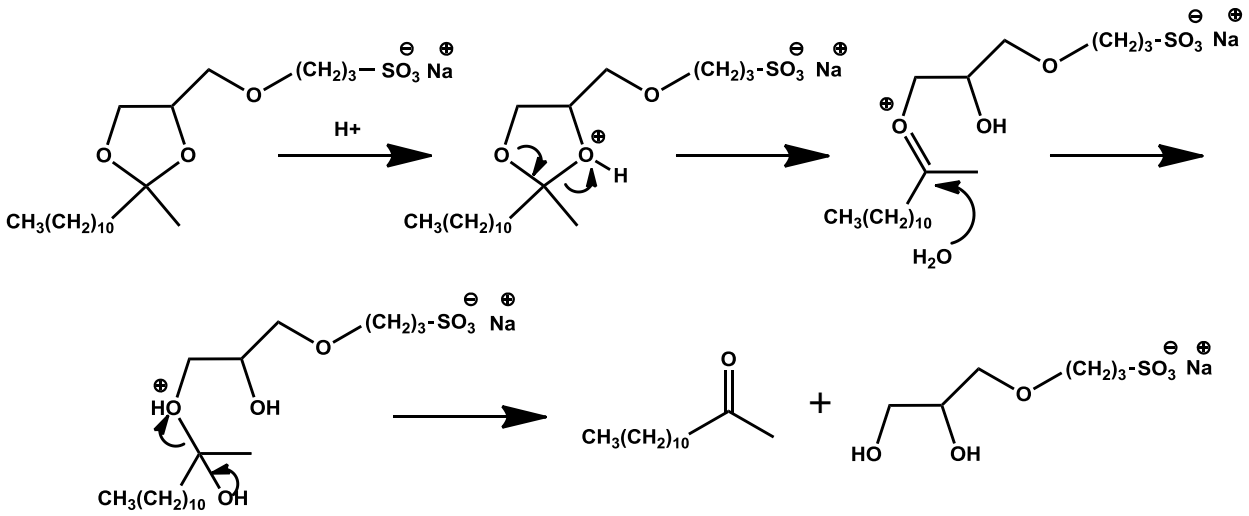
In tissue proteomics studies using flash-frozen tissue samples, sample contamination with blood proteins is an intractable problem (1). Herein, we employed a sequential tissue extraction method that was designed to minimize blood protein contamination. In this sequential extraction method, tissue samples were first washed thoroughly using phosphate buffered saline (PBS), to remove blood covering the tissue, and then homogenized in HEPES buffer (100 μL buffer : 1 mg tissue). After centrifugation, the supernatants were saved as the Pre-extract, which presumably contained highly abundant soluble proteins such as albumin and other blood proteins. The pellets from the pre-extraction were then homogenized in HEPES buffer a second time without surfactant (Control) or with either MaSDeS or SDS at 0.1%, 0.2%, or 0.5% surfactant concentration (**Figure 1E-F**).

We next sought to determine whether or not the enhanced solubilization of proteins observed in extracts containing MaSDeS was due to increased solubilization of blood proteins. A total of 1,423 individual proteins were identified in the Pre-extract, Control, and MaSDeS extracts from heart tissue with 74% of the proteins identified in the Pre-extract also being detected in either the Control extract, MaSDeS extract, or both (**Supplemental Figure S5A** and **Supplemental Table S4**). Of the 173 proteins (26%) that were uniquely identified in the Pre-extract, 80% (138) belonged to the plasma/blood proteome according to the plasma proteome database (http://www.plasmaproteomedatabase.org/index.html) (**Supplemental Table S4**),

suggesting that the Pre-extract contains a high abundance of soluble blood proteins. Indeed, Western blot (**Supplemental Figure S5B**) and label-free quantitative MS analyses (**Supplemental Figure S5C-F**) confirmed the significant reduction in the amount of common blood proteins such as albumin, which accounts for approximately half of the total protein in plasma (2), in both the Control and MaSDeS-containing heart tissue extracts. In contrast, no differences were observed in the abundances of the cytosolic house-keeping proteins GAPDH, HSP90, and tubulin between extracts (**Supplementary Figure S5G-I**). The reason could be that during the pre-extraction, the tissue samples were only briefly homogenized, which might not be sufficient to break apart the cells. Collectively, the data shows that the Pre-extract contains a greater abundance of blood proteins (compared to Control and MaSDeS extracts) and that the removal of these proteins in the pre-extraction allows for a greater number of proteins to be identified in the subsequent extraction.

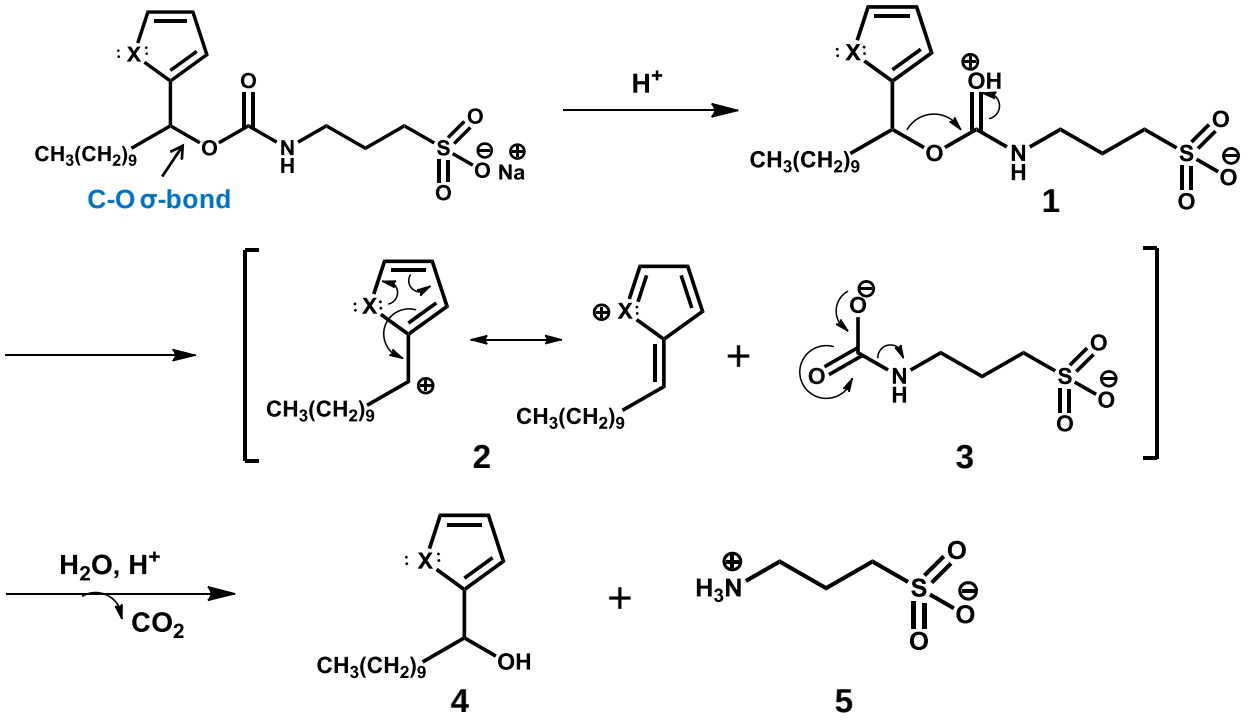
Supplemental Scheme 1. Comparison of the degradation mechanisms of RapiGest (RG), ProteaseMax (PM) and MaSDeS under acidic conditions.

A. RG



RG has a ketal functional group, which is susceptible to acid hydrolysis. At pH <5, this functional group is easily fragmented into a hydrophilic anionic molecule and a hydrophobic molecule ($CH_3(CH_2)_{10}COCH_3$) as described previously by Yu et al. (3).

B. PM (X=O) vs MaSDeS (X=S)



PM and MaSDeS have similar molecular structures that differ only in their labile groups, which contain (i) a furan ring (X=O) and (ii) a thiophene ring (X=S), respectively. Generally, at pH <2 (acidic conditions), the carbonyl oxygen atom would be protonated and lead to the degradation of molecule 1. The furanyl group (X=O) is a better electron-donating group than the thiophene group (X=S) and, thus, cleavage of the C-O σ-bond (see an arrow flow in molecule 1) will occur more quickly in PM. Subsequently, intermediate 3 undergoes decarboxylation [loss of carbon dioxide (CO_2)]. This provides a strong driving force for the degradation of the surfactant into CO_2 and a zwitterionic species (3-ammoniopropane-1-sulfonate) (molecule 5)(4). Intermediate 2 then undergoes the addition of water, producing an alcohol 4. Therefore, it is expected that PM will be more easily degraded into CO_2, R-OH (molecule 4), and a small zwitterionic species (molecule 5) due to its furan ring, compared to MaSDeS, which contains a thiophene ring.

References:

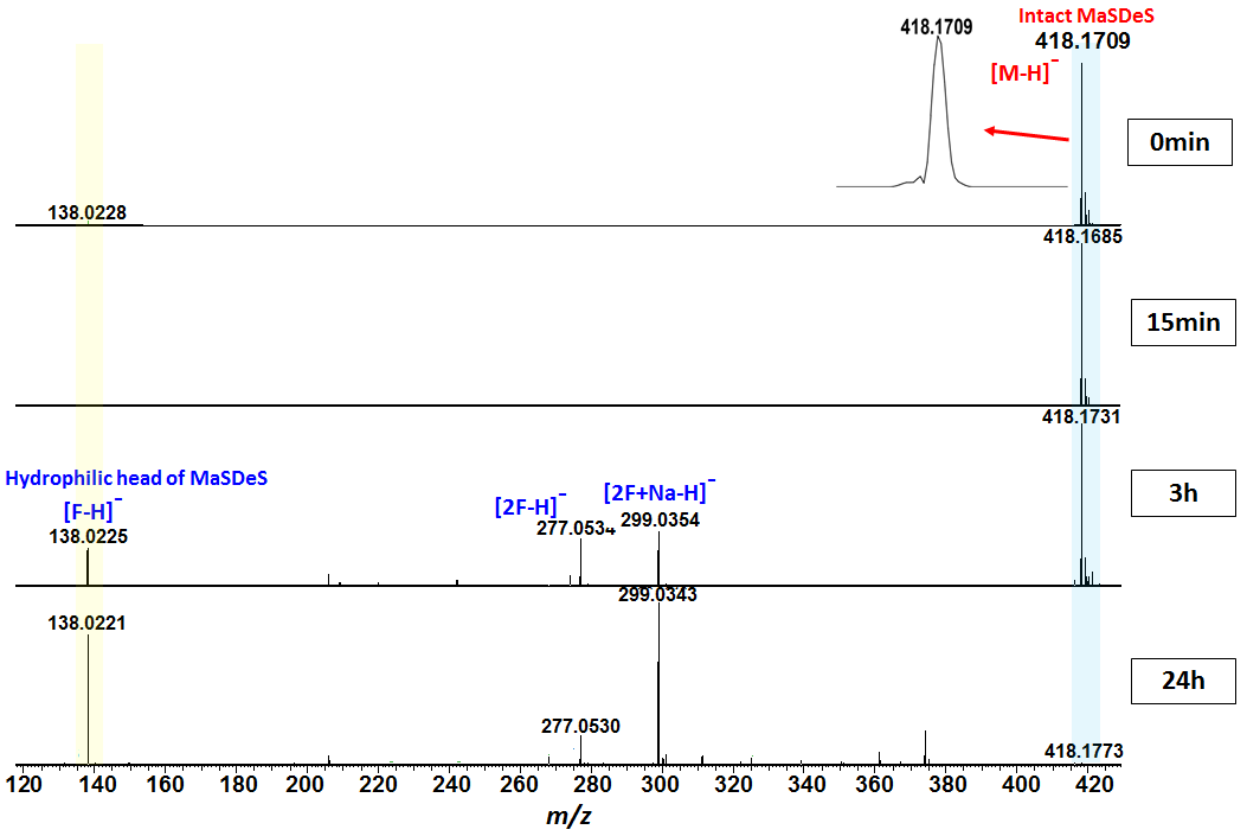
1. Kim, M. S., Pinto, S. M., Getnet, D., Nirujogi, R. S., Manda, S. S., Chaerkady, R., Madugundu, A. K., Kelkar, D. S., Isserlin, R., Jain, S., Thomas, J. K., Muthusamy, B., Leal-Rojas, P., Kumar, P., Sahasrabuddhe, N. A., Balakrishnan, L., Advani, J., George, B., Renuse, S., Selvan, L. D., Patil, A. H., Nanjappa, V., Radhakrishnan, A., Prasad, S., Subbannayya, T., Raju, R., Kumar, M., Sreenivasamurthy, S. K., Marimuthu, A., Sathe, G. J., Chavan, S., Datta, K. K., Subbannayya, Y., Sahu, A., Yelamanchi, S. D., Jayaram, S., Rajagopalan, P., Sharma, J., Murthy, K. R., Syed, N., Goel, R., Khan, A. A., Ahmad, S., Dey, G., Mudgal, K., Chatterjee, A., Huang, T. C., Zhong, J., Wu, X., Shaw, P. G., Freed, D., Zahari, M. S., Mukherjee, K. K., Shankar, S., Mahadevan, A., Lam, H., Mitchell, C. J., Shankar, S. K., Satishchandra, P., Schroeder, J. T., Sirdeshmukh, R., Maitra, A., Leach, S. D., Drake, C. G., Halushka, M. K., Prasad, T. S., Hruban, R. H., Kerr, C. L., Bader, G. D., Iacobuzio-Donahue, C. A., Gowda, H., and Pandey, A. A draft map of the human proteome. *Nature* 2014, 509, 575-581.

2. Rifai, N., Gillette, M. A., and Carr, S. A. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol* 2006, *24*, 971-983.

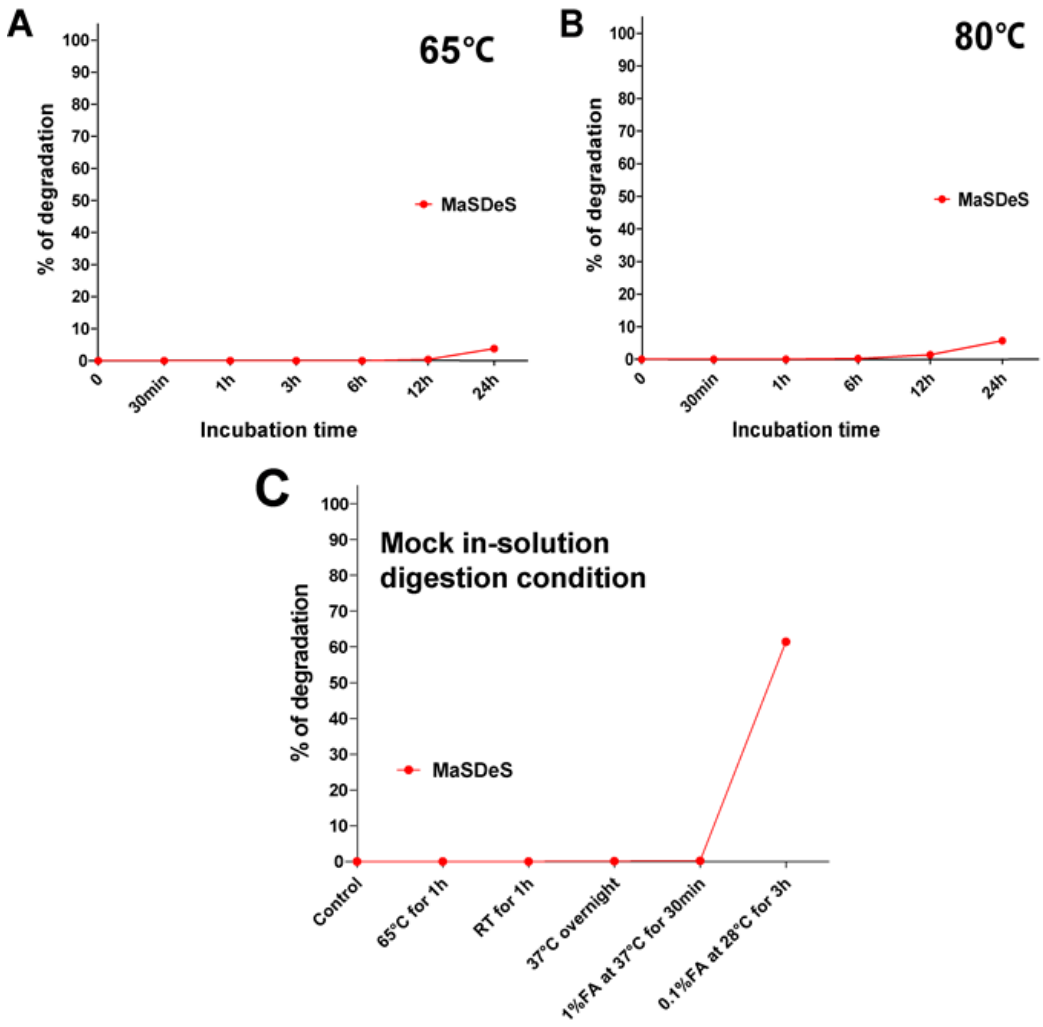
3. Yu, Y. Q., Gilar, M., Lee, P. J., Bouvier, E. S., and Gebler, J. C. Enzyme-friendly, mass spectrometry-compatible surfactant for in-solution enzymatic digestion of proteins. *Anal. Chem.* 2003, 75, 6023-6028.

4. Saveliev, S. V., Woodroofe, C. C., Sabat, G., Adams, C. M., Klaubert, D., Wood, K., and Urh, M. Mass Spectrometry Compatible Surfactant for Optimized In-Gel Protein Digestion. *Anal. Chem.* 2012, 85, 907-914.

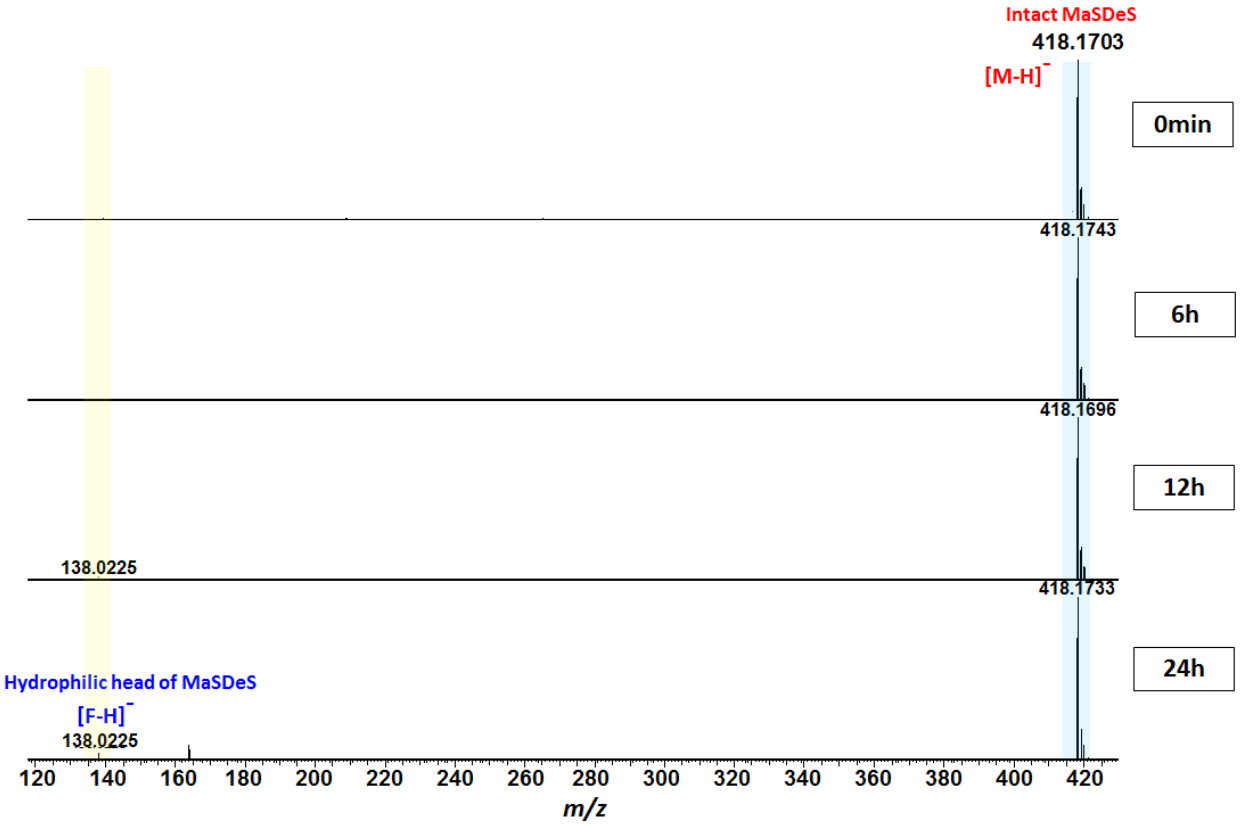
Supplemental Figures



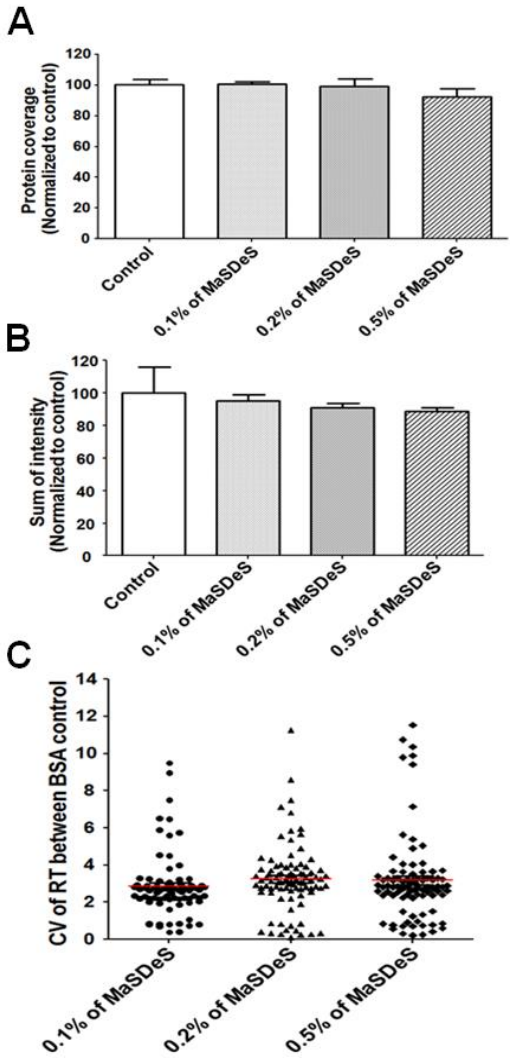
Supplemental Figure S1. High-resolution Fourier transform ion cyclotron resonance (FT-ICR) MS analysis of MaSDeS degradation in 10% formic acid (FA). The negative ion mode MS data was collected at 0 min, 15 min, 30 min, 1 h, 3 h, 6 h, 12 h, and 24 h to investigate the degradation rate of MaSDeS in 10% FA at 37 °C. The MS data (at 0min, 15min, 3h, and 24h) shows that the hydrophilic head of MaSDeS (*m/z* 138.02) started to appear at 15 min and then gradually increased in intensity until the intact surfactant (*m/z* 418.17) was completely depleted by the 24 h time-point. $[M-H]^-$, molecular ion of the intact MaSDeS; $[F-H]^-$, molecular ion of the hydrophilic head of MaSDeS; $[2F-H]^-$, and $[2F-Na]^-$ molecular ion of the dimer form of the hydrophilic head of MaSDeS and sodiated forms, respectively.



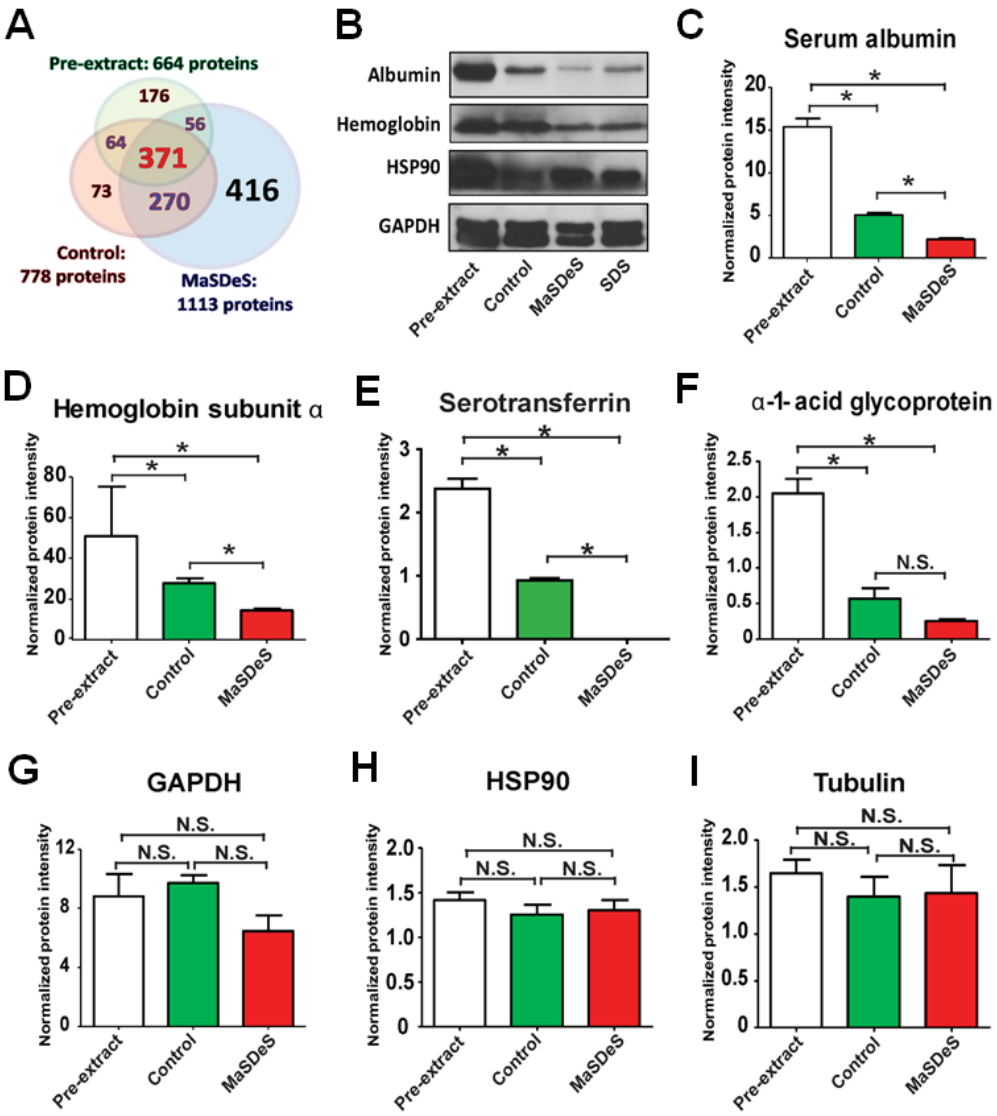
Supplemental Figure S2. Assessment of MaSDeS thermo-stability. Assessment of MaSDeS degradation over a 24 h time period at (**A**) 65 °C and (**B**) 80 °C. (**C**) Assessment of MaSDeS degradation under mock in-solution digestion conditions. Control: before any procedures; 65°C for 1 h: to mimic the dithiothreitol (DTT) reaction step; RT for 1h: to mimic the iodoacetamide (IAA) reaction step; 37 °C overnight: to mimic the trypsin digestion step; 1% FA at 37°C for 30 min: to mimic stopping of the trypsin digestion; 0.1% FA at 28 °C for 3h: digested samples were kept at 4°C for a day and then incubated at 28 °C for 3 h in 0.1% FA to mimic LC-MS/MS run conditions.



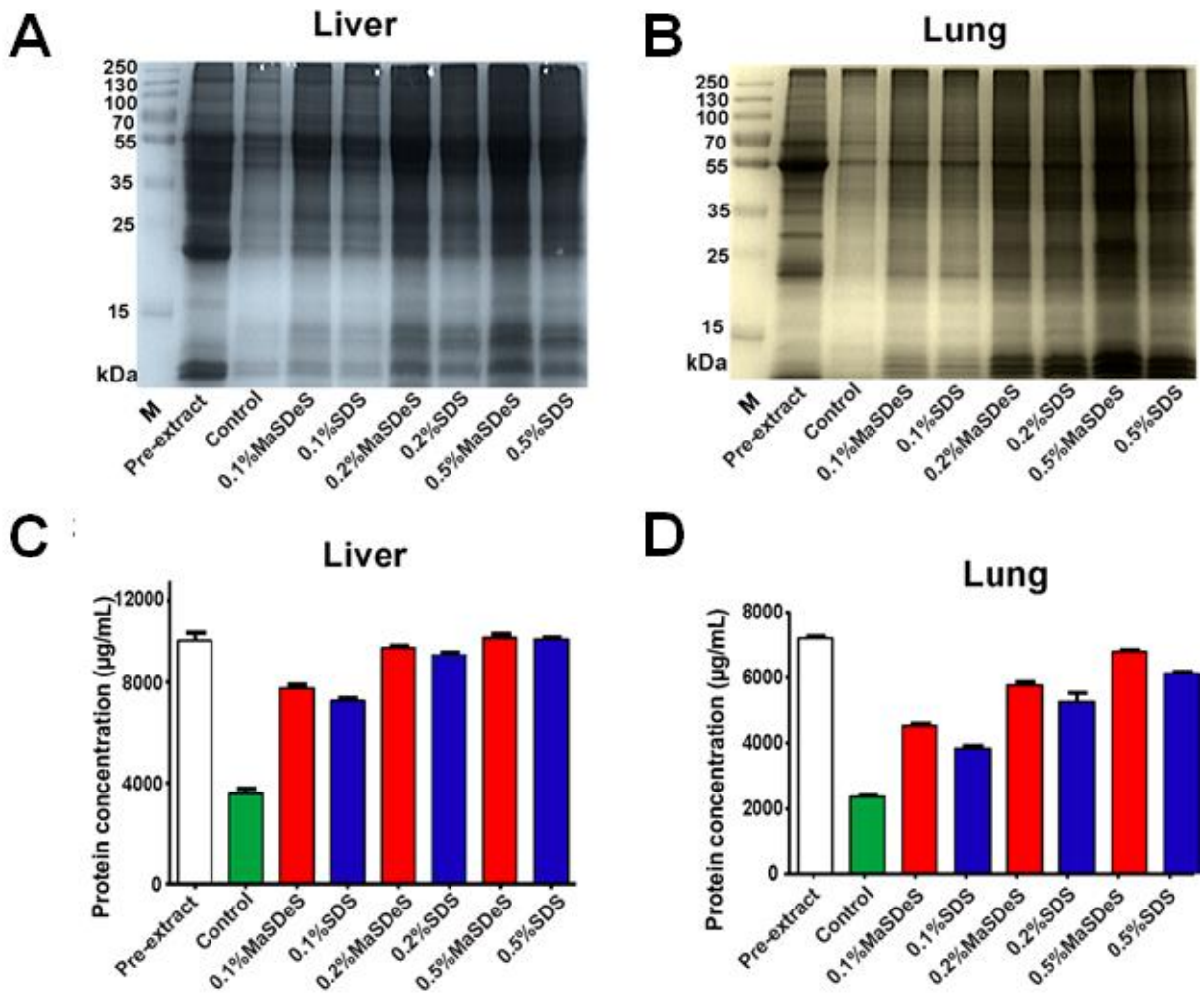
Supplemental Figure S3. MS evaluation of MaSDeS stability at 65 °C. The FT-ICR MS spectra for MaSDeS (*m/z* 418.17) show almost no degradation of the surfactant over the 24 h incubation period at 65 °C.



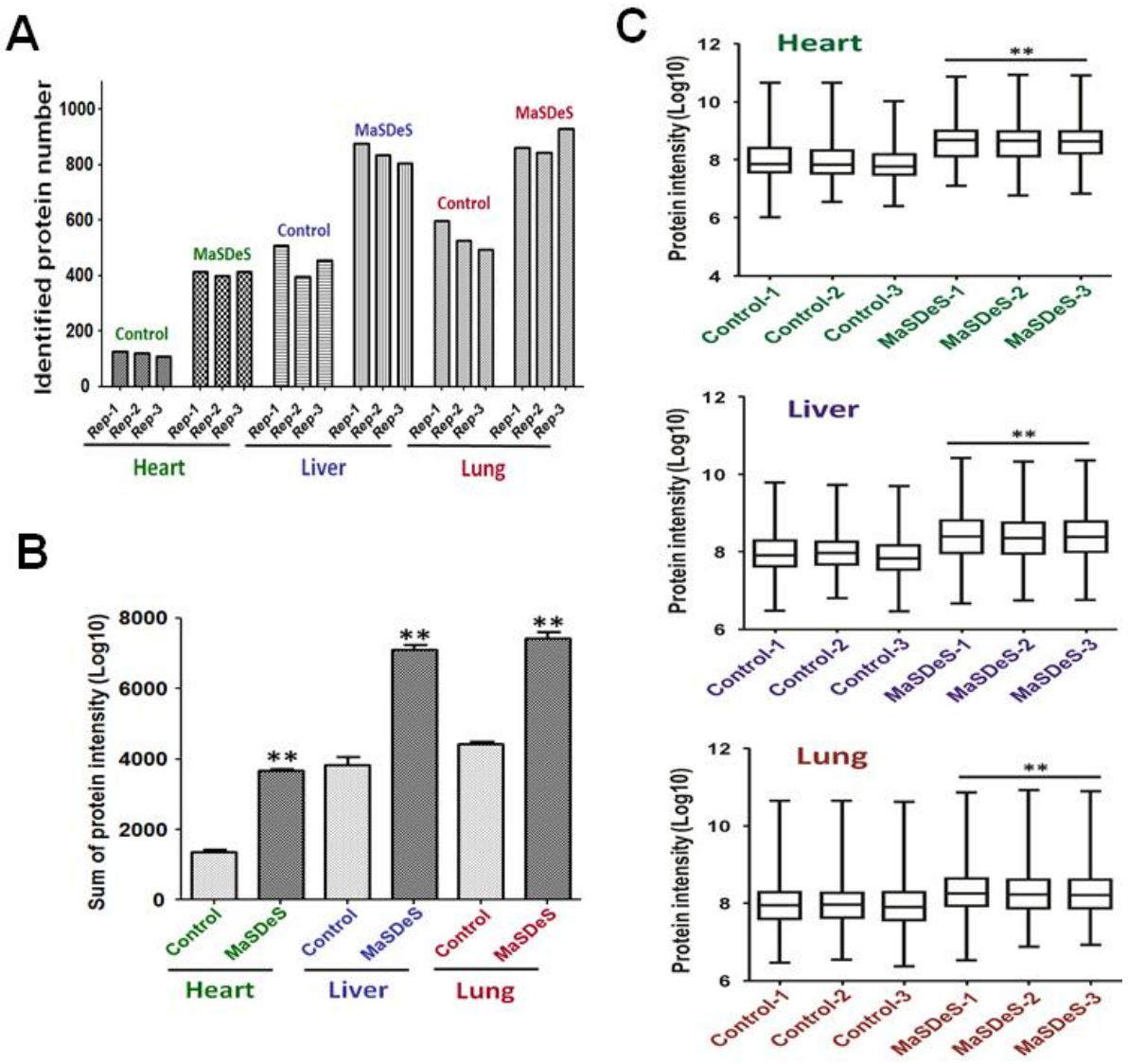
Supplemental Figure S4. Evaluation of the influence of MaSDeS on enzymatic digestion and MS analysis. MaSDeS was spiked into samples containing bovine serum albumin (BSA), at final concentrations of 0.1%, 0.2%, and 0.5% surfactant, followed by in-solution digestion and MS analysis. The experiments were run in triplicate. Three indicators were used to determine whether MaSDeS could influence digestion efficiency and MS analysis: (**A**) protein coverage, (**B**) the sum of the peptide intensity, and (**C**) the coefficient of variation (CV) of peptide retention time (RT) between BSA without surfactant and BSA with either 0.1%, 0.2%, or 0.5% MaSDeS, respectively. Each dot indicates a peptide. There were no significant differences between the results obtained for BSA without MaSDeS and BSA with 0.1%, 0.2%, or 0.5% MaSDeS for all three indicators, suggesting that MaSDeS has minimal influence on enzymatic digestion and MS analysis.



Supplemental Figure S5. The sequential tissue extraction method reduces blood protein contamination in the second heart tissue extract. (**A**) A Venn diagram showing the number of protein identified in Pre-extract, Control, and 0.2% MaSDeS-containing extracts as well as the overlap between these extracts. (**B**) Western blot data and (**C-F**) quantitative MS results illustrating the significant reduction in the abundances of blood proteins in extracts generated using buffer that contained either no (Control) or 0.2% MaSDeS. (**G-I**) In contrast, the abundance of cytosolic proteins was not significantly different between extracts. *, $p<0.05$. N.S., not significant.

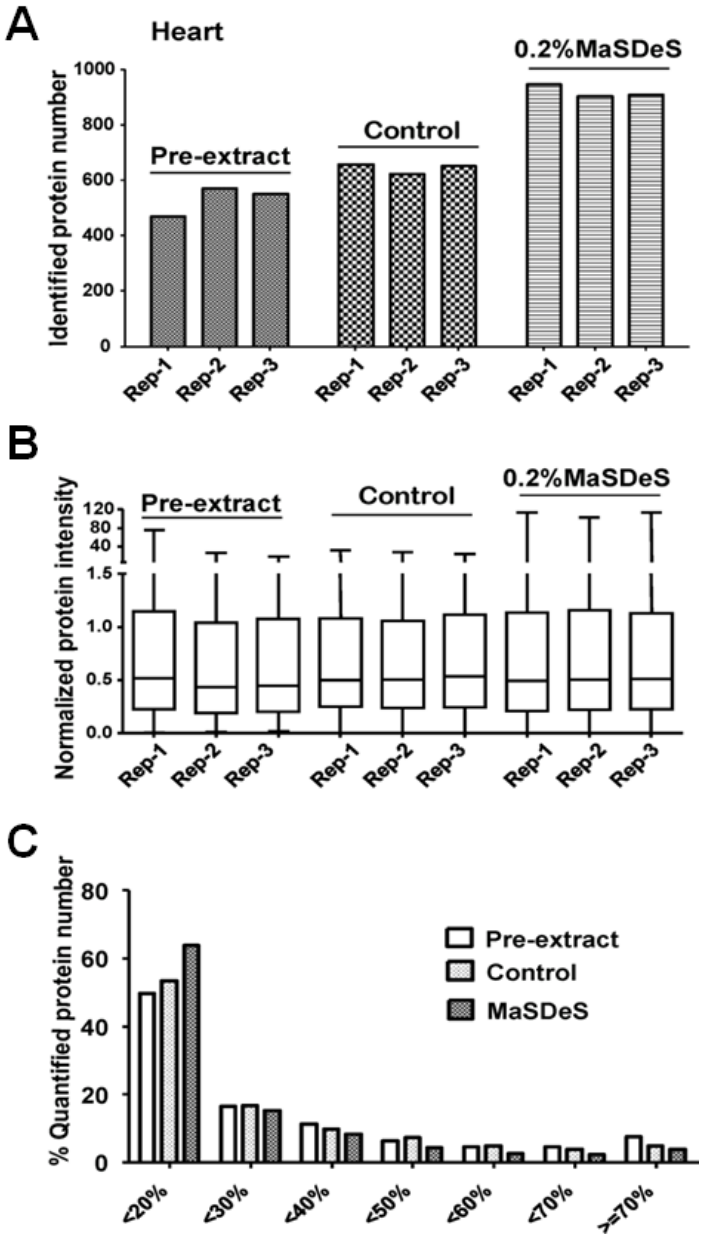


Supplemental Figure S6. SDS-polyacrylamide gel electrophoresis (PAGE) analysis of sequential extracts from (**A**) liver and (**B**) lung using buffer with 0.1%, 0.2%, or 0.5% of MaSDeS or SDS, respectively. Bradford protein assay results of extractions from (**C**) liver and (**D**) lung are presented. An equal volume of each extract was evaluated using these methods. Both SDS-PAGE and Bradford protein assay results show that protein solubilization in extracts containing MaSDeS is comparable to that obtained with SDS.



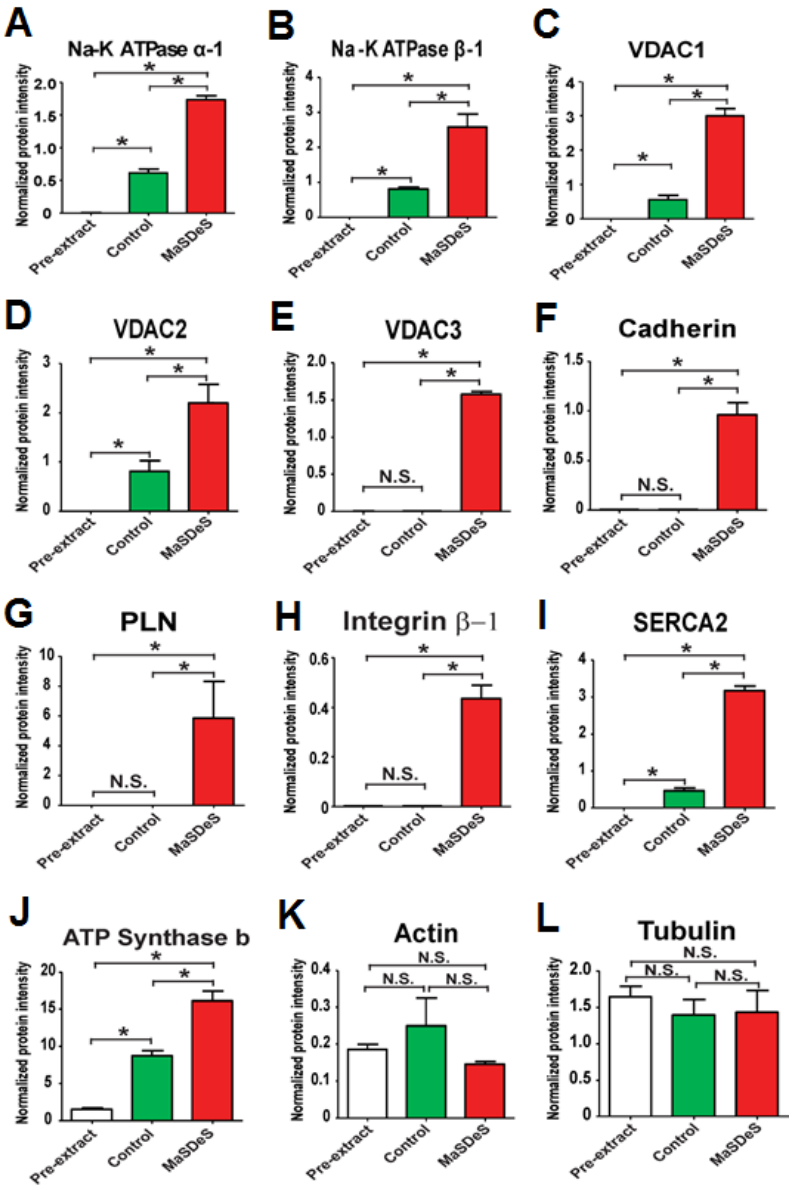
Supplemental Figure S7. The effect of MaSDeS on protein identification. Equal volumes of extracts from heart, liver, and lung tissues generated with buffer without or with 0.2% MaSDeS were analyzed using MS to assess the number of proteins identified in commonly used tissues. Analyzing the same volume (1.5 μL) of extract from extracts without and with MaSDeS allowed us to assess the ability of the surfactant to solubilize proteins from different tissues and to identify them using MS. Each sample was run in triplicate. (**A**) The number of proteins identified in tissue extracts with or without MaSDeS. The addition of MaSDeS to the extraction buffer increased the number of proteins identified in the extract. (**B**) Sum of protein intensities in each extract from heart, liver, and lung tissues. For the quantification of equal volume injection, the area under the curve of each peptide was calculated using Proteome Discoverer and the average

of the three most abundant distinct peptides served as the final protein intensity. All of the given protein intensities are presented in Log10 form. In the experiment of equal volume injection, the quantified abundances were not normalized by any factor in order to show the differences in protein abundance between Control and MaSDeS-containing tissue extracts when an equal volume of extract was analyzed. Under equal volume loading, the extracts with added MaSDeS showed a higher sum of intensities in comparison to Control extracts. These results show that the inclusion of MaSDeS in the extraction buffer increased protein identifications as well as the overall abundance of protein in the extract. (**C**) Under equal volume loading, the MaSDeS-containing extracts showed higher protein intensity distributions in each run compared to Control extracts. The protein intensities were simply presented by log10 transformations of the raw abundances obtained from Proteome Discoverer. * means the sample has a statistically significant difference from control sample. *, $p < 0.05$. **, $p < 0.01$.

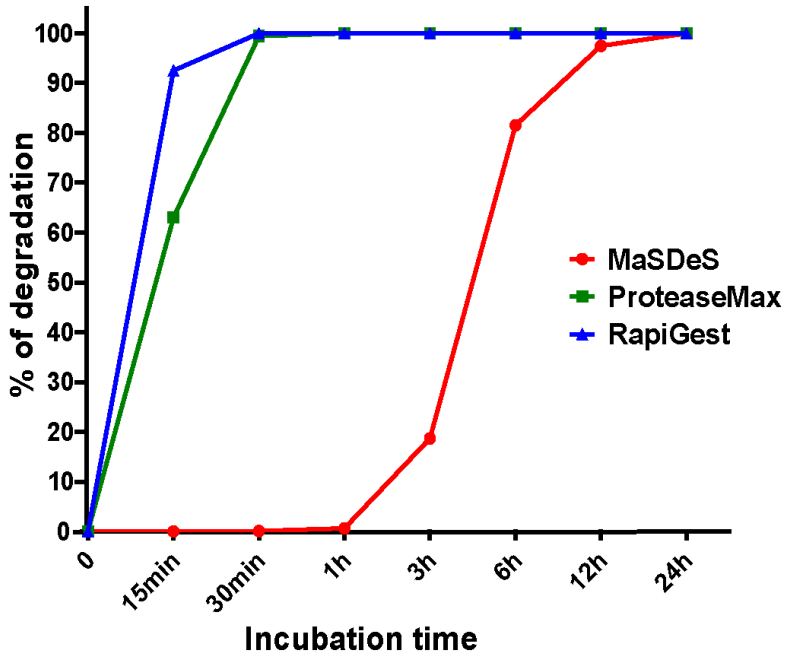


Supplemental Figure S8. MS analysis of equal protein amounts (1 µg) from pig heart tissue extracts (Pre-extract, Control, and 0.2% MaSDeS) shows improved protein identification in extracts generated with buffer containing MaSDeS. Loading equal amounts of protein allowed us to determine if, for the same amount of total protein, additional proteins were solubilized in extracts containing 0.2% MaSDeS versus those without surfactant, as well as to compare protein abundances between samples. Pig heart tissue was used and the experiments were run in triplicate. (**A**) The number of proteins identified in each extract. The purpose of the first extraction is to reduce blood protein contamination. After the first extraction, protein

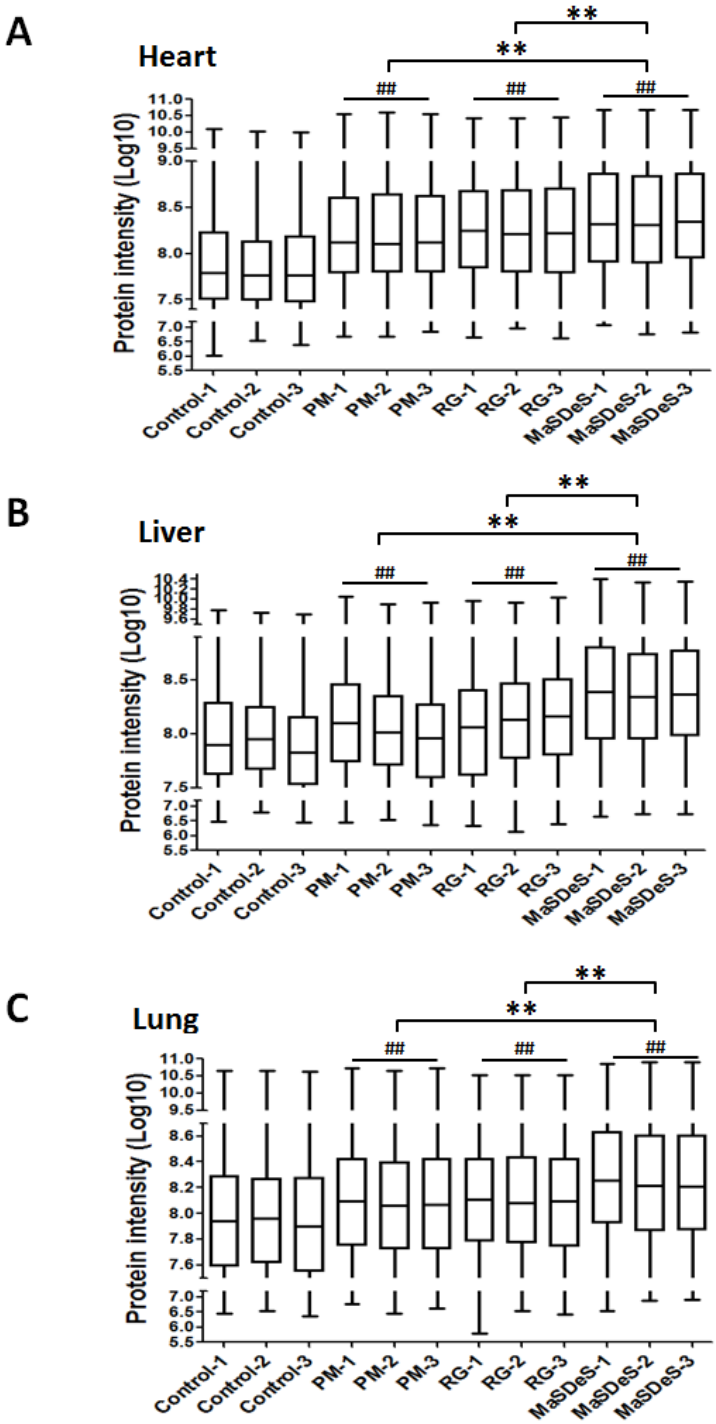
identification in the subsequent extraction was improved, likely as a consequence of reduced contamination from blood proteins. (**B**) In the experiment of equal amount injection, the peptide intensities were normalized by certain factors to minimize the variation between runs. Normalized peptide intensity was used to adjust for median signal variances from run-to-run. All peptides corresponding to their respective proteins in each run were calculated for their geometric mean, which provided the protein's final abundance. The intensity distribution for each protein showed there was no statistical difference between the extracts after normalization. The lowest protein intensity divided by 10 was assigned to unidentified proteins. (**C**) The coefficient of variation (CV) within a group showed reliable data for quantification after normalization. Approximately, 80% of protein CVs were under 30%.



Supplemental Figure S9. Quantitative MS results showing that the use of MaSDeS significantly improved the detection of membrane proteins in heart tissue extracts. The MS data set presented here is the same as the data in **Figure 2** with the inclusion of the Pre-extract for comparison. (**A-J**) Membrane proteins were most abundant in extracts with added MaSDeS. In contrast, the abundances of cytosolic proteins such as (**K-L**) actin and tubulin were not influenced. *, $p<0.05$; N.S., not significant.



Supplemental Figure S10. MaSDeS degrades slowly under acidic conditions (10 % FA in 25 mM ammonium bicarbonate (ABC) buffer) at 37 °C, in contrast to the rapid degradation observed for the two commercially available acid-labile surfactants, RapiGest (RG) and ProteaseMAX (PM).



Supplemental Figure S11. Buffer containing 0.2% of PM, RG, or MaSDeS was used to homogenize pig heart, liver, and lung tissue. An equal volume of each sample (1.5 μL) was injected into LC-MS/MS to evaluate the capability of each surfactant for protein identification. Each sample was run in triplicate. The intensity distribution showed that under equal volume

loading, the MaSDeS-containing extracts showed higher protein intensity distributions in each run compared to the Control, PM, and RG extracts. The protein intensities were simply presented by Log10 transformations of the raw abundances obtained from Proteome Discoverer. * means the sample has a statistically significant difference from MaSDeS sample. *, $p < 0.05$. **, $p < 0.01$. # means the sample has a statistically significant difference from Control sample. #, $p < 0.05$. ##, $p < 0.01$.